



MINISTÉRIO DA EDUCAÇÃO
FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA
CAMPUS DE JI-PARANÁ
DEPARTAMENTO DE MATEMÁTICA E ESTATÍSTICA



GILBERTO FRANCISCO DE PAULA JUNIOR

**AJUSTAMENTO DE MODELOS DE PROBABILIDADE PARA VARIÁVEIS DE
PRECIPITAÇÃO NA CIDADE DE CACOCAL/RO**

Ji-Paraná/RO
2016

GILBERTO FRANCISCO DE PAULA JUNIOR

**AJUSTAMENTO DE MODELOS DE PROBABILIDADE PARA VARIÁVEIS DE
PRECIPITAÇÃO NA CIDADE DE CACOCAL/RO**

Trabalho de Conclusão de Curso apresentado na
Fundação Universidade Federal de Rondônia –
Campus de Ji-Paraná, como parte das exigências
do Curso de Graduação em Bacharelado em
Estatística, para a obtenção do título de Bacharel
em Estatística.

Orientador: Dilson Henrique Ramos Evangelista

Dados Internacionais de Catalogação na Publicação
Biblioteca Setorial – UNIR/Campus de Ji-Paraná

Paula Júnior, Gilberto Francisco de
P324a Ajustamento de modelos de probabilidade para variáveis de
2016 precipitação na cidade de Cacoal/RO / Gilberto Francisco de Paula
Júnior; orientador, Dilson Henrique Ramos Evangelista. – Ji-Paraná,
2016
29 f. : 30 cm

Trabalho de conclusão do curso de Estatística. – Universidade
Federal de Rondônia, 2016
Inclui referências

1. Probabilidade e estatística. 2. Estatística matemática.
I. Evangelista, Dilson Henrique Ramos. II. Universidade Federal de
Rondônia. III. Título

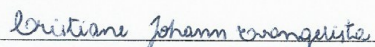
CDU 519.2





ATA DA SESSÃO DE APRESENTAÇÃO E DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO


Aos 30 (trinta) dias do mês de maio de 2016, realizou-se na Sala Laboratório de Estatística 2, no Campus de Ji-Paraná, a Sessão de Apresentação e Defesa do Trabalho de Conclusão de Curso (TCC), com o título “**Ajustamento de modelos de probabilidade para variáveis de precipitação na cidade de Cacoal/RO**”, apresentado pelo acadêmico **Gilberto Francisco de Paula Júnior**. Os trabalhos foram instalados às 09:00 horas pelo presidente da Banca Examinadora, aprovada pelo Departamento e constituída por: Prof^º. **Dr^ª Roziane Sobreira dos Santos** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística), Prof^º. **Msc. Cristiane Johann Evangelista** (Universidade Estadual Paulista) e Prof^º **Dr. Dilson Henrique Ramos Evangelista** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística), sendo este o orientador do acadêmico. A Banca Examinadora, tendo decidido aceitar o Trabalho de Conclusão de Curso, após a apresentação passou à arguição pública do acadêmico. Encerrando os trabalhos de arguição às 09:54 horas. A Banca Examinadora deu parecer final **Aprovado**, com a nota 95 (**noventa e cinco**), resultado da média aritmética das notas individuais atribuídas pelos membros da Banca Examinadora. Proclamado o resultado final pelo presidente da Banca, foram encerrados os trabalhos. Para constar, lavrou-se a presente Ata, que é assinada pelos membros da Banca Examinadora e pelo acadêmico.

Ji-Paraná, 30 de maio de 2016.


Msc. Cristiane Johann Evangelista
Prof^º .Membro da Banca


Dr.ª. Roziane Sobreira dos Santos
Prof^º. Membro da Banca


Dr. Dilson Henrique Ramos Evangelista
Prof^º Orientador e Pres. Banca


Gilberto Francisco de Paula Júnior
Acadêmico

Dedicatória

“Dedico este trabalho, primeiramente a Deus por me dar forças, em seguida a minha família: ao meu pai Gilberto e a minha mãe Idenir, e meus dois queridos irmãos Rogerio e Edina”.

Epígrafe

“Eu odeio toda a atenção. Eu acho que é melhor deixar meu trabalho falar por mim”.

Shigeru Miyamoto

AGRADECIMENTOS

A Deus e meu senhor Jesus Cristo pela força e saúde para superar as dificuldades do dia a dia.

Aos meus familiares, em especial meu pai Gilberto e minha mãe Idenir e aos meus dois queridos irmãos Edina e Rogerio que me acompanharam em minha caminhada durante o curso.

As minhas duas queridas Avós maternas e paternas Mirtes Apolinária (Im Memorian) e Nadir de Paula (Im Memorian).

Ao querido Professor orientador Dilson Henrique Ramos Evangelista que me orientou, e me ajudou na ideia do tema para que fosse possível a execução deste trabalho.

Ao INMET (Instituto Nacional de Meteorologia) pelos dados e informações cedidas.

A UNIR (Fundação Universidade Federal de Rondônia) pela oportunidade de cursar um curso superior.

RESUMO

JUNIOR, Gilberto Francisco de Paula, Discente, Universidade Federal de Rondônia, Ji-Paraná. **AJUSTAMENTO DE MODELOS DE PROBABILIDADE PARA VARIÁVEIS DE PRECIPITAÇÃO NA CIDADE DE CACOCAL/RO.** 2016. Orientador: Dilson Henrique Ramos Evangelista.

O presente trabalho teve como objetivo, realizar ajustamento das distribuições de probabilidade em que, foram utilizados os dados de precipitação da cidade de Cacoal/RO fornecidos pelo INMET. Os dados estudados são referente aos meses de janeiro, fevereiro, março, outubro, novembro e dezembro entre 2010 a 2014. Foram analisadas cinco distribuições densidade de probabilidade (Exponencial, Log-normal, Gama, Weibull e Normal). Em cada distribuição foi testada a aderência usando os testes não paramétrico de Qui-Quadrado e Kolmogorov-Smirnov, a um nível de 5% de significância. O mês de fevereiro apresentou maior nível de precipitação com média superior de 12,94mm e desvio padrão de 16,40mm. Nos resultados obtidos do ajustamento as distribuições Weibull, Normal, Log-normal e Gama apresentaram resultados satisfatórios nos testes de aderência, os dados analisados só não se adequaram a distribuição exponencial nos dois testes realizados.

Palavras-chave: Ajustamento, Teste de Aderência, Probabilidade, Precipitação.

ABSTRACT

JUNIOR, Gilberto Francisco de Paula, Student, Universidade Federal de Rondônia, Ji-Paraná. **ADJUSTMENT BY PROBABILITY MODELS FOR VARIABLE PRECIPITATION IN THE CITY OF CACOCAL / RO.** 2016. Advisor: Dilson Henrique Evangelista Ramos.

This study aimed, perform adjustment of probability distributions in which the data of precipitation of the city of Cacoal / RO provided by INMET were used. The data studied are for the months of January, February, March, October, November and December from 2010 to 2014 five distributions probability density were analyzed (Exponential, Log-Normal, Gamma, Weibull and Normal). In each distribution was tested for adhesion using the nonparametric tests of Chi-square and Kolmogorov-Smirnov, at a 5% level of significance. The month of February showed a higher level of precipitation average higher 12,94mm and standard deviation of 16,40mm. The results of adjusting the Weibull, Normal, Log-normal range and showed satisfactory results in adhesion tests, the data analyzed just did not fit the exponential distribution in both tests.

Keywords: Adjustment, Adhesion Test, Probability, rainfall.

LISTA DE FIGURA

Figura 1: Totais de precipitação da região de Cacoal nos meses de janeiro, fevereiro, março, outubro novembro e dezembro entre 2010 a 2014.....	25
Figura 2 - Gráficos CDF de função de distribuição acumulada empírica das distribuições Normal, Log-Normal, Gama e Weibull.....	27

LISTA DE EQUAÇÕES

- 1 Função Densidade de Probabilidade da Distribuição Normal
- 2 Função Densidade de Probabilidade da Distribuição Log-Normal
- 3 Função Densidade de Probabilidade da Weibull
- 4 Função Acumulada da Distribuição Weibull
- 5 Ajuste da distribuição Weibull Pelo Método da Máxima Verossimilhança
- 6 Ajuste da distribuição Weibull Pelo Método da Máxima Verossimilhança
- 7 Função Densidade de Probabilidade da Exponencial
- 8 Função Acumulada da Exponencial
- 9 Função Densidade de Probabilidade da Gama
- 10 Função de Distribuição Gama
- 11 Estimativa da Gama Usando o Método da Máxima Verossimilhança
- 12 Estimativa da Gama Usando o Método da Máxima Verossimilhança
- 13 Estimativa da Gama Usando o Método da Máxima Verossimilhança
- 14 Estimativa da Gama Usando o Método da Máxima Verossimilhança
- 15 Estimativa da Gama Usando o Método da Máxima Verossimilhança
- 16 Estimativa da Gama Usando o Método da Máxima Verossimilhança
- 17 Estatística do Teste do Qui-Quadrado
- 18 Regra de decisão do Qui-Quadrado
- 19 Estatística de Teste do Kolmogorov-Smirnov

LISTA DE TABELAS

Tabela 1 - Análise descritiva dos meses no período de 2010 a 2014.....	25
Tabela 2 - Ajustamento das funções de densidades estudadas, com base nos testes de Qui-Quadrado e Kolmogorov-Smirnov.....	26

LISTA DE ABREVEATURAS E SIMBOLOS

FDP	Função Densidade de Probabilidade
IBGE	Instituto Brasileiro de Geografia e Estatística
INMET	Instituto Nacional de Meteorologia
KS	Kolmogorov-Smirnov

SUMÁRIO

1. INTRODUÇÃO	15
1.2 OBJETIVOS	16
1.2.1 GERAIS.....	16
1.2.2 ESPECIFICOS	17
2. METODOLOGIA.....	17
3. REFERENCIAL TEORICO	18
3.1 ESTATÍSTICAS DESCRITIVAS.....	18
3.2 DISTRIBUIÇÕES DE PROBABILIDADES	18
3.2.1 DISTRIBUIÇÃO NORMAL.....	19
3.2.2 DISTRIBUIÇÃO LOG-NORMAL	19
3.2.3 DISTRIBUIÇÃO WEIBULL	19
3.2.4 DISTRIBUIÇÃO EXPONENCIAL	20
3.2.3 DISTRIBUIÇÃO GAMA.....	21
3.3 TESTES DE ADERÊNCIA.....	22
3.3.1 TESTE DE ADERÊNCIA DO QUI-QUADRADO (χ^2)	23
3.3.2 TESTE DE ADERÊNCIA DE KOLMOGOROV-SMIRNOV (KS).....	24
4. RESULTADOS E DISCUSSÕES	25
4.1 ANÁLISE DESCRITIVA	25
4.2 AJUSTES DAS DISTRIBUIÇÕES	25
5. CONCLUSÃO.....	28
REFERÊNCIAS	29
APÊNDICE – ROTINA DA ANÁLISE DO AJUSTAMENTO	31

1. INTRODUÇÃO

A Amazônia está localizada ao norte da América do Sul e ocupa uma área total de mais de 6,5 milhões de quilômetros quadrados. Aproximadamente 85% da Amazônia ficam em território brasileiro, e abrange aproximadamente 61% da área do país. Administrativamente, a Amazônia brasileira é composta pelos estados do Acre, Amapá, Amazonas, Pará, Rondônia, Roraima, além de parte dos estados de Mato Grosso, Tocantins e Maranhão (BRASIL, 2009).

Com sua vasta extensão territorial, a Amazônia apresenta diferentes características climáticas e meteorológicas em suas sub-regiões, e de superfície, em termos de vegetação e solos com diferentes impactos de atividades humanas (DIAS, 2006). Nos últimos anos, a comunidade científica vem se interessando cada vez mais em estudar o comportamento dos elementos meteorológicos para caracterizar a variabilidade temporal dos elementos climáticos, que é indispensável para quantificar adequadamente os efeitos ocasionados, isto é um dos grandes problemas para estudos na região amazônica.

Quantificar a manifestação dinâmica ocasionada por processos naturais através de função de distribuição de probabilidade e seus parâmetros têm sido uma das grandes aplicações na comunidade científica; Vários trabalhos, sendo eles o de Neto et al. (2005), Cruciani et al. (2002), Catalunha et al. (2002) e outros apresentam aplicação de técnica da função de distribuição de probabilidade (FDP) em elementos climatológicos.

De acordo com Catalunha et al. (2002) “O uso de funções densidade de probabilidade (FDP) está inteiramente ligado à natureza dos dados a que ela se relaciona, algumas têm boa aptidão de estimação para pequeno número de dados, outras requerem grandes séries de observações”. No caso de dados meteorológicos, requer grande número de elementos na amostra para validar os estudos estatísticos e seu comportamento climatológico ao longo dos anos. Séries climatológicas apresentam uma variação temporal muito alta, conforme Back (2001) oscilação climática é a flutuação na qual a variável tende a mover-se gradativamente e de forma cordial entre sucessivos máximos e mínimos.

As técnicas de FDP são aplicadas segundo as características da variável em estudo, tendo uma complexibilidade em conformidade com cada variável na amostra, pois em determinada situação podem assumir vários números de estimativa de parâmetros (DUAN ET AL., 1998).

As FDP com os seus pressupostos satisfeitos trazem muitas informações sobre a população da variável estudada, principalmente na capacidade de estimar os dados observados, de acordo com cada parâmetro, e a capacidade de estimação é medida com teste de aderência (CATALUNHA ET AL., 2002).

No Brasil, a região norte apresenta umas das maiores precipitações do país, os quais em algumas áreas podem atingir valores superiores a 3.000 mm (NIMER, 1979). A principal influência é por conta da Floresta Amazônica, na qual é produzida um grande volume de vapor de água devido à evapotranspiração da floresta.

A precipitação é um dos fenômenos definido como qualquer deposição de água em forma líquida ou sólida sobre a superfície da terra, em forma de gelo, neve, granizo ou chuva. Para a medição da precipitação, todos os valores são expressos em milímetros (mm) de água líquida equivalente para o intervalo de tempo precedente (ou em polegadas). Um milímetro de chuva corresponde a 1 litro por metro quadrado de água sobre a superfície (METEOBLUE, 2015).

Silva et al. (2007) realizaram um trabalho de uma análise de distribuição de probabilidade na cidade de Santa Maria/RS, na qual os autores utilizaram cinco funções de distribuição de probabilidade (Gama, Weibull, Normal, Log-normal e Exponencial). Os objetivos dos autores eram analisar a distribuição diária de precipitação e verificar os números de dias com chuva para determinar a variação de precipitação diária, durante os meses do ano. Segundo os autores as funções que melhor se ajustaram aos dados foram Gama e Weibull, pois não houveram diferenças significativas entre estas duas funções. Os meses de junho e julho as chuva são mais frequente, porém é semelhante nos demais meses do ano.

Para este trabalho o objetivo foi analisar e ajustar os cinco modelos de distribuições de probabilidades na série histórica de precipitação nos meses de janeiro, fevereiro, março, outubro, novembro e dezembro entre 2010 a 2014 no município de Cacoal/RO.

1.2 OBJETIVOS

1.2.1 GERAIS

Analisar a série de dados de precipitação da cidade de Cacoal/RO, selecionar dentre as cinco distribuições (Exponencial, Gama, Log-normal, Normal e Weibull) a que melhor se ajusta.

1.2.2 ESPECÍFICOS

- a) Revisão sobre os cinco modelos de distribuição estudados para o ajuste;
- b) Análise descritiva da serie temporal dos dados (média, variância e desvio padrão);
- c) Ajustar os modelos de distribuição de probabilidade (Exponencial, Gama, Log-normal, Normal e Weibull), analisar a aderência dos cinco modelos estudados.

2. METODOLOGIA

Cacoal é um dos 52 municípios do estado de Rondônia, localizado na porção centro leste do estado, na microrregião do leste rondoniense, a uma latitude 11°26'19 sul e a uma longitude 61°26'50" oeste, estando numa altitude de 200 metros. Sua população segundo o IBGE/2010 é de 78.574 habitantes, sendo considerada a quarta maior cidade do estado, possui uma área de 3.793 km² representando 1,6% do estado onde seu território tem como limite os municípios de: Presidente Médici ao noroeste, Espigão d'Oeste ao leste, Castanheiras e ministro Andreazza ao Oeste, Pimenta Bueno ao sudoeste e Rolim de Moura ao sul.

Segundo Leão (2008) a região de Cacoal se caracteriza por elevadas precipitações cujo total compensa a estação seca que permite a existência de floresta, possui um período chuvoso que ocorre entre os meses de outubro a abril, e o período seco em junho, julho e agosto, e os meses de maio e setembro são de transição. Cacoal tem um clima do tipo tropical que chove muito mais no verão que no inverno, com uma temperatura média em torno de 24°C.

Para o desenvolvimento deste trabalho foi realizada revisões bibliográficas sobre ajustamento de probabilidade, envolvendo livros, revistas científicas nacionais e internacionais, dissertações, teses e sites da internet que abordam questões relacionadas à precipitação, caracterização da região de estudo, ajustamento e testes de aderências.

A partir da obtenção da série temporal dos dados de precipitação fornecida pelo INMET (Instituto Nacional de Meteorologia), da estação meteorológica automática de Cacoal/RO com as coordenadas 11°26'S, 61°26'W e 210m, foram empregadas técnicas estatísticas de descrição dos dados para avaliar o comportamento das séries de precipitação do local de estudo, e ajustar relativamente a uma distribuição probabilística adequada, através dos testes de aderência do Qui-Quadrado e Kolmogorov-Smirnov ao nível de 5% de

significância. Os dados obtidos referem-se a precipitações pluviais (mm) diárias dos meses de janeiro, fevereiro, março, outubro, novembro e dezembro entre 2010 a 2014.

Foram utilizados nas análises os seguintes softwares: Microsoft Office Excel 2010 com o complemento do XLSTAT para a elaboração das estatísticas descritivas, e o R (Core TEAM, 2006) com os pacotes MASS e riskDistributions para a elaboração dos ajustes.

3. REFERENCIAL TEÓRICO

3.1 ESTATÍSTICAS DESCRITIVAS

Inicialmente, antes dos ajustes de probabilidades, foi realizada a análise das estatísticas descritivas; como por exemplo, as medidas de posição como as médias, mediana e moda, e medidas de dispersão como a variância, coeficiente de variação, amplitude e desvio padrão. Estas medidas permitem identificar os meses de maiores precipitações na região, e detectar as variabilidades e compreender o comportamento da série dos dados. A estatística descritiva é uma das etapas iniciais de uma análise, onde permite resumir os dados para que se possam ser expresso por meio de tabelas e gráficos. Neste trabalho foram utilizadas as três principais medidas: média, variância, e desvio padrão para a elaboração das análises descritivas dos meses citados.

3.2 DISTRIBUIÇÕES DE PROBABILIDADES

As séries temporais de variáveis hidrológicas como as precipitações quando são observadas ao longo do tempo, apresentam variações sazonais, porém estas variações não são absolutamente regulares.

Fisch (1999) desenvolveu um trabalho sobre a distribuição de precipitação na área da cidade de Taubaté/SP no vale do rio Paraíba, em seu artigo foi realizada uma análise numa série temporal de 15 anos, o objetivo do autor era a determinação do início e final do período chuvoso da região. O início do período começa na primeira semana de outubro e se estende até a primeira semana de maio, os meses chuvosos são dezembro, janeiro e fevereiro, já no período de inverno são os meses secos junho, julho e agosto.

A variável hidrológica de precipitação esta sempre associada a uma distribuição de probabilidade de ocorrência. A utilização ou não de uma distribuição de probabilidade reside na capacidade da mesma estimar os dados observados com base em seus parâmetros, esta capacidade é medida com base na aplicação de testes de aderência (ALMEIDA, 1995).

3.2.1 DISTRIBUIÇÃO NORMAL

Segundo Morettin & Bussab (2004) a distribuição normal é uma das mais importantes distribuições para as variáveis aleatórias contínuas. A distribuição serve para muitos problemas da vida real, mas também aparece em muitas investigações teóricas nas quais muitas técnicas estatísticas como análise de variância, regressão, e alguns testes de hipóteses, assumem ou exigem uma normalidade dos dados. Uma variável aleatória $X \in \mathbb{R}$ tem uma distribuição normal se sua função de densidade de probabilidade for da forma de:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right], \quad -\infty < x < +\infty \quad (1)$$

Onde:

μ : média populacional

σ : desvio padrão populacional.

3.2.2 DISTRIBUIÇÃO LOG-NORMAL

A Log-Normal é uma distribuição que se assemelha muito com uma normal, pois é uma generalização de uma distribuição normal, porém trabalha-se com o logaritmo nos dados. Segundo DeGroot & Schervish (1975) a distribuição Log-Normal possui a seguinte função densidade de probabilidade a dois parâmetros:

$$f(x) = \frac{1}{(2\pi)^{1/2}\sigma x} \exp \left(-\frac{(\log(x) - \mu)^2}{2\sigma^2} \right), \quad (2)$$

Onde μ e σ são os parâmetros de distribuição, média e desvio padrão respectivamente, e x é o valor da variável aleatória.

3.2.3 DISTRIBUIÇÃO WEIBULL

A distribuição Weibull é muito utilizada em análise hidrológica para eventos extremos. A distribuição foi proposta por Waloddi Weibull em 1954, no qual lançou um artigo descrevendo a distribuição em detalhes e propondo diversas aplicações, a distribuição pode ser aplicada em diversas áreas do conhecimento, como por exemplo, em variáveis hidrológicas. Segundo Catalunha et al. (2002) a função de probabilidade da Weibull é apresentada de diversas formas, em alguns trabalhos é apresentada como:

$$f(x) = \frac{\gamma}{\beta} \left(\frac{x - \alpha}{\beta} \right)^{\gamma-1} \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^\gamma \right] \text{ para } x \geq \alpha \quad (3)$$

Onde, $f(x) = 0$ para outros intervalos, x é o valor da variável aleatória e $\alpha \geq 0, \beta \geq 0$ e $\gamma > 0$ os parâmetros da distribuição. No caso de Weibull de três parâmetros é dada pela seguinte distribuição acumulada:

$$F(x) = \int_0^\infty f(x) dx = 1 - \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^\gamma \right] \quad (4)$$

As razões de a distribuição ser muito aplicada em diversas situações é a íntima relação entre seus parâmetros e comportamentos da distribuição. O parâmetro α é chamado de parâmetro de posição, pois controla a posição da curva sobre o eixo das abscissas. O parâmetro de escala β controla as dimensões que a curva assume, dada uma forma constante. O parâmetro γ , é chamado parâmetro de forma (CATALUNHA, 2000).

Um dos principais métodos de ajuste da distribuição Weibull, é o da máxima verossimilhança (COUTO, 1980), este método consiste em determinar os valores de γ e β pelas suas equações fundamentais. Nota que, entretanto β é função de γ , e que a solução desse tipo de sistema (Equação 5) é obtido resolvendo a equação por um processo iterativo, como por exemplo o método da bissecção sendo β encontrado, a partir de γ aproximado e introduzido na Equação 6.

$$\frac{\sum_{i=1}^n x_i^{\hat{\gamma}} \ln(x_i)}{\sum_{i=1}^n x_i} - \frac{\sum_{i=1}^n \ln(x_i)}{n} - \frac{1}{\hat{\gamma}} = 0 \quad (5)$$

$$\hat{\beta} = \left(\frac{\sum_{i=1}^n \ln(x_i^{\hat{\gamma}})}{n} \right)^{\frac{1}{\hat{\gamma}}} \quad (6)$$

3.2.4) DISTRIBUIÇÃO EXPONENCIAL

Por definição a distribuição exponencial se caracteriza por ter uma função de taxa de falha constante. É uma das distribuições mais simples em termos matemáticos, geralmente aplicada a dados com forte assimetria, quando estão apresentando uma função de “J” invertido. Além deste fato, é uma das distribuições que possui inúmeras aplicações em diversas áreas do conhecimento, e em particular em variáveis hidrológicas. Uma variável

aleatória X tem distribuição exponencial, se sua função de densidade de probabilidade for descrita por (NAGHETTINI & PINTO, 2007):

$$f(x) = \lambda \exp(-\lambda x) \text{ para } x \geq 0 \quad (7)$$

Onde sua função acumulada é dada por:

$$F(x) = 1 - \exp(-\lambda x) \quad (8)$$

Onde que o único parâmetro da distribuição (λ) visto é estimado pelo inverso da média.

3.2.5) DISTRIBUIÇÃO GAMA

A gama é uma distribuição contínua, é uma das distribuições mais gerais, pois diversas distribuições são caso particular dela como, por exemplo, a distribuição exponencial, qui-quadrado e entre outras. Esta distribuição é frequentemente aplicada na determinação da probabilidade da precipitação para durações de dias, semanas, meses, e até anos.

Segundo Catalunha et al. (2002) se uma variável aleatória continua X com ($0 \leq x \leq \infty$) com distribuição gama de parâmetros $\alpha > 0$ e $\beta > 0$, se sua função de densidade de probabilidade é:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad 0 < x < \infty \quad (9)$$

Sendo a função de distribuição dada abaixo:

$$F(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x u^{\alpha-1} e^{-\frac{u}{\beta}} du \quad (10)$$

Onde:

x : quantidade de chuva em mm;

$F(x)$: probabilidade de ocorrer uma quantidade de precipitação igual ou inferior a x ;

α : parâmetro de forma da variável aleatória x ;

β : é o parâmetro de escala da variável aleatória de x ;

Uma das formas de estimar os parâmetros da distribuição gama é o método de máxima verossimilhança. THOM (1958), usando o método da máxima verossimilhança, derivou as equações para estimativa dos parâmetros da distribuição gama através da seguinte equação quadrática:

$$12 A \alpha^2 - 6\alpha - 1 = 0 \quad (11)$$

$$\hat{\alpha} = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad (12)$$

A estimativa para o parâmetro de escala foi feita por:

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \quad (13)$$

Em que:

$$A = \ln \bar{x} - x_g \quad (14)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (15)$$

$$x_g = \frac{1}{n} \sum_{i=1}^n \ln(x_i) \quad (16)$$

3.3) TESTES DE ADERÊNCIA

Para se ajustar uma série temporal dos dados, é necessário verificar se existe uma adequação entre as probabilidades ou frequências. Os testes de aderência são instrumentos da estatística que auxiliam na tomada de decisão da adequação ou inadequação de um modelo numa certa amostra, por exemplo, permite verificar se uma variável aleatória contínua segue uma distribuição segundo um determinado modelo ou não. Os objetivos dos testes de aderência em si, é testar a adequabilidade de um modelo probabilístico a um conjunto de dados (NAGHETTINI & PINTO, 2007).

As próximas sessões abordaram com mais detalhes os dois testes de aderências que foram usados neste trabalho, os testes de aderência de Qui-Quadrado (χ^2) e Kolmogorov-Smirnov (KS).

3.3.1) TESTE DE ADERÊNCIA DO QUI-QUADRADO (χ^2)

O teste de aderência do Qui-Quadrado (χ^2) proposto por Karl Pearson em 1900 é um teste que consiste em comparar os dados obtidos experimentalmente com os dados esperados de acordo com uma lei probabilística. Segundo Naghettini & Pinto (2007) a estatística de teste do Qui-Quadrado é:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (17)$$

Onde:

O_i = frequência observada, ou seja, o total de indivíduos observados na categoria, $i = 1, 2, \dots, n$.

E_i = frequência esperada, ou seja, total de indivíduos esperados na categoria, $i = 1, 2, \dots, n$.

Na regra de decisão pode ser baseada no valor de P, ou seja:

$$P = P(\chi^2 \geq \chi_{obs}^2) \quad (18)$$

Em que χ_{obs}^2 é o valor calculado a partir dos dados usando a equação 17. Se para α fixo, obtermos $P \leq \alpha$, rejeita-se a hipótese H_0 .

Hipóteses:

H_0 : o comportamento pressupõe um “bom” ajustamento na presente distribuição de probabilidade, ou seja, os valores estão próximos dos observados.

H_1 : o comportamento não pressupõe um “bom” ajustamento na presente distribuição de probabilidade, ou seja, os valores não estão próximos dos observados.

Em geral, pode-se dizer que a estatística Qui-Quadrado nada mais é do que a distância quadrática entre os valores observados e esperados, ou seja, quanto maior o valor dessa estatística, maior é a evidência de que os dados não seguem a presente distribuição de probabilidade fornecida.

3.3.2) TESTE DE ADERÊNCIA DE KOLMOGOROV-SMIRNOV (KS)

O teste de aderência de Kolmogorov-Smirnov é um teste não paramétrico, cuja sua estatística de teste tem como base a diferença máxima entre as funções de probabilidades acumuladas, empíricas e teóricas, de variáveis aleatórias contínuas. O teste em si não é aplicado em variáveis aleatórias discretas, pois é um teste para dados contínuos (NAGHETTINI & PINTO 2007).

A ideia do teste de Komogorov-Smirnov é avaliar a função de distribuição acumulada $F(x) = P(X \leq x)$. Suponha-se que foi selecionada uma função de distribuição desconhecida $F(x)$. O teste ira avaliar se $F(x) = F_n(x)$ para todo x comparando $F_n(x)$ com $F(x)$ que é obtido da amostra.

Estatística de teste:

$$D_n = \sup |F(x) - F_n(x)| \quad (19)$$

Onde:

D_n : distância máxima vertical entre os gráficos de $F(x)$ e $F_n(x)$ sobre a amplitude dos possíveis valores de x ;

$F(x)$: representa a função de distribuição acumulada assumida para os dados;

$F_n(x)$: representa a função de distribuição empírica dos dados.

A regra de decisão é rejeitar H_0 se as estatísticas de teste D_n é maior que o quantil $1-\alpha$ para o teste Kolmogorov-Smirnov. Caso contrário, não se rejeita a hipótese H_0 .

Hipóteses:

H_0 : O comportamento probabilístico da variável aleatória, em questão pode ser modelado pela seguinte distribuição x de probabilidade.

H_1 : O comportamento probabilístico da variável aleatória, em questão não pode ser modelado pela seguinte distribuição x de probabilidade.

4. RESULTADOS E DISCUSSÕES

4.1) ANÁLISE DESCRITIVA

A Tabela 1 abaixo apresenta os resultados da análise descritiva por mês, pode-se observar que os meses de maiores precipitações são de Dezembro, Janeiro, Fevereiro e Março por apresentar totais de precipitações superiores a 1.400,00mm.

Tabela 1: Análise descritiva dos meses no período de 2010 a 2014

Estatísticas descritivas				
Meses	Total (mm)	Média	Variância	Desvio Padrão
Janeiro	1.737,90	11,21	254,36	15,95
Fevereiro	1.824,81	12,94	268,96	16,40
Março	1.598,65	10,31	340,53	18,45
Outubro	872,26	5,63	172,34	13,13
Novembro	1.044,20	6,96	183,61	13,55
Dezembro	1.482,09	9,56	223,45	14,95

Fonte: Próprio autor

Em relação à média e ao desvio padrão os meses de fevereiro apresentaram maiores médias com 12,94mm, e os meses de março apresentaram maiores variações nos dados, com variância de 340,53mm e desvio padrão de 18,4mm.

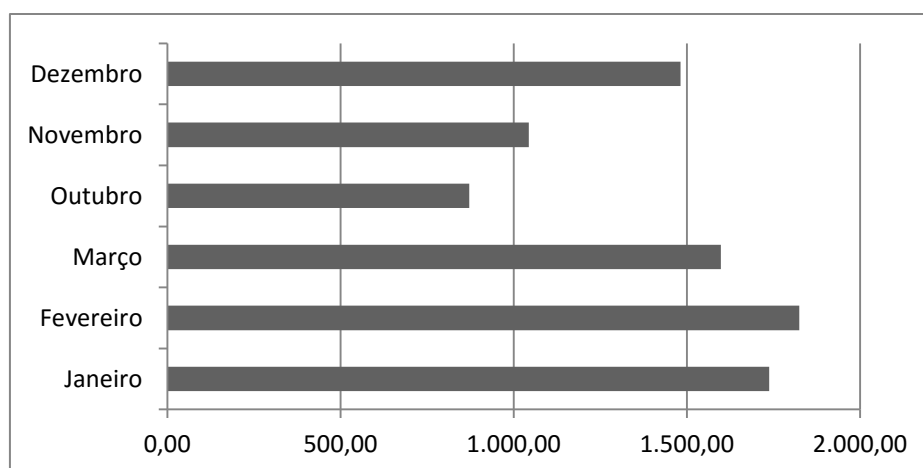


Figura 2: Totais de precipitação da região de Cacoal nos meses de janeiro, fevereiro, março, outubro novembro e dezembro entre 2010 a 2014.

Fonte: Próprio Autor

4.2) AJUSTES DAS DISTRIBUIÇÕES

Para o ajustamento das distribuições de probabilidade, cada distribuição foi testada a sua aderência nos testes de Kolmogorov-Smirnov e Qui-Quadrado foi utilizado $\alpha = 5\%$ de significância para a aceitação dos testes.

Tabela 2: Ajustamento das funções de densidades estudadas, com base nos testes de Qui-Quadrado e Kolmogorov-Smirnov.

Distribuição	Qui-Quadrado	Kolmogorov-Smirnov
	p-valor	p-valor
Gama	0,011*	0,321
Weibull	0,054	0,600
Log-Normal	0,008*	0,254
Normal	0,045*	0,670
Exponencial	<0,0001*	0,014*

Fonte: Próprio autor **Observação:** * significativo para $\alpha = 0,05$ de significância.

Observa-se na Tabela 2 que a distribuição exponencial foi rejeitada nos dois testes de aderências realizados. As distribuições Gama, Weibull, Log-Normal e Normal foram aceitas e apresentaram resultados satisfatórios no teste de Kolmogorov-Smirnov. Na realização do teste do Qui-quadrado somente a distribuição Weibull teve aceitação. Em geral pode-se interpretar que a Normal e Weibull são as duas distribuições que melhor se ajustaram no presente conjunto de dados com índice de significância de 5%. A Normal apesar de não ter tido aceitação no teste de qui-quadrado, foi a que obteve maior p-valor que foi de 0,670 do teste de Kolmogorov-Smirnov, entretanto a Weibull teve aceitação em ambos os testes, sendo assim foi considerado que tanto a Normal e a Weibull são as duas distribuições que apresentaram melhores resultados de um ajustamento, pois a diferença entre os dois valores do p-valor do teste de Qui-Quadrado e Kolmogorov-Smirnov são muito próximas.

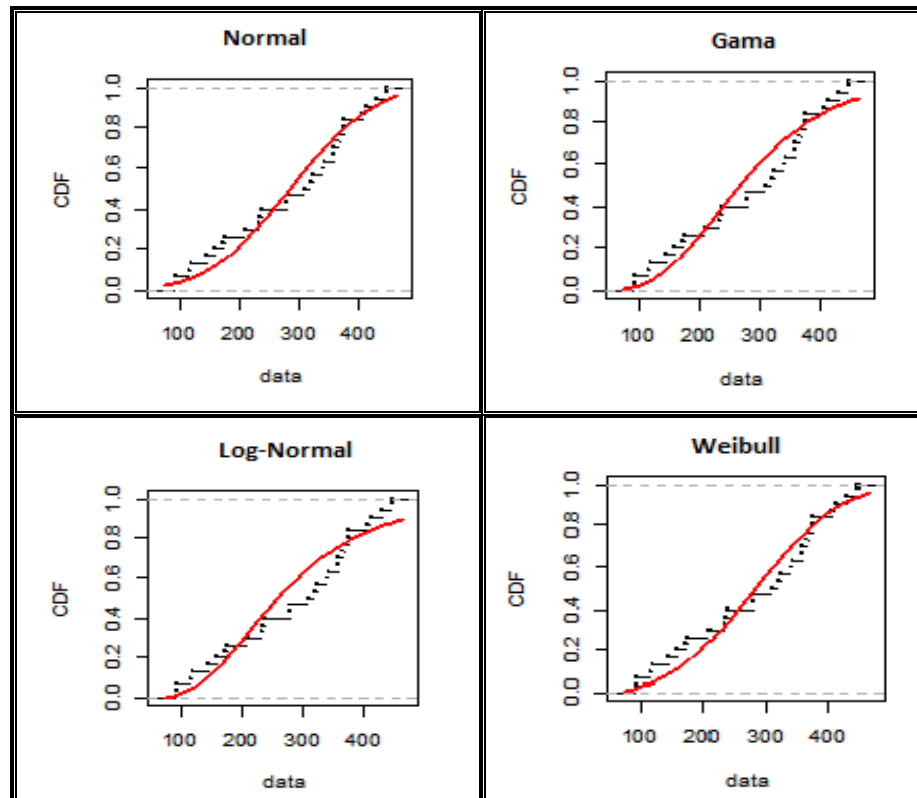


Figura 2: Gráficos CDF de função de distribuição acumulada empírica das distribuições Normal, Log-Normal, Gama e Weibull.

Fonte: Próprio autor

A figura 2 acima apresenta os gráficos CDF das distribuições Normal, Log-Normal, Gama e Weibull. O CDF (Cumulative Distribution Function) ou Função de distribuição cumulativa é um gráfico que é utilizado para avaliar um ajuste de uma distribuição para seus dados, estimar percentis e comparar diferentes distribuições. Na figura 2 foram incluídos apenas os gráficos das distribuições que se ajustaram pelos testes de Kolmogorov-Smirnov e Qui-Quadrado. Pode-se observar que em geral as presentes distribuições citadas acima apresentam um bom ajuste nos dados.

5. CONCLUSÃO

Os meses de maiores precipitações são de Dezembro, Janeiro, Fevereiro e Março por apresentar totais de precipitações superiores a 1.400,00mm. Os meses de fevereiro apresentaram as maiores médias com 12,94mm e os meses de março apresentaram maior variação e desvio padrão dos dados com variância de 340,53mm e desvio padrão de 18,4mm.

Conforme nos resultados obtidos do ajustamento as distribuições Weibull, Normal, Log-Normal e Gama foram as que apresentaram resultados satisfatórios nos testes de aderências do Qui-Quadrado e Kolmogorov-Smirnov, entretanto a distribuição Normal e a Weibull são as duas distribuições que apresentam melhores resultados de um ajustamento, pois os valores de p-valor são maiores que das outras distribuições estudadas, e a diferença entre os dois valores do p-valor das duas distribuições no teste do Qui-Quadrado e Kolmogorov-Smirnov são muito próximas. A distribuição exponencial por sua vez foi rejeitada nos dois testes de aderências realizados.

REFERÊNCIAS

- ALMEIDA, R. M. B. **Características climatológicas do regime de chuva em Minas Gerais**, Dissertação (Mestrado em Engenharia Agrícola) – Curso de pós-graduação em Engenharia Agrícola, Universidade Federal de Viçosa, Viçosa, 64p, 1995.
- BACK, A. J. Seleção de Distribuição de probabilidade para chuvas diárias extremas no estado de Santa Catarina. **Revista Brasileira de Meteorologia**, v. 16, n. 2, pp. 211-222, 2001.
- CATALUNHA, M. J. **Avaliação do ajustamento de funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais**, Dissertação de Mestrado, Mestrado em Meteorologia Agrícola, Universidade Federal de Viçosa, MG, 87p, 2000.
- CATALUNHA, M. J.; SEDIYAMA, G. C.; LEAL, B. G.; SOARES, C. P. B.; RIBEIRO, A. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 10, n. 1, p. 153-162, 2002.
- COUTO, H.T.Z. **Distribuição de diâmetros em plantações de Pinus caribaea Morelet var caribaea**. 1980. 79p. Tese (Livre Docência) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP, 1980.
- CRUCIANI, D.E.; MACHADO, R.E.; SENTELHAS, P. C. Modelos da distribuição temporal de chuvas intensas em Piracicaba, SP. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, n. 1, p. 76-82, 2002.
- DEGROOT, H.M; SCHERVISH, M.J. **Probability and statistics**. 1975.
- DIAS, R. **Gestão Ambiental: Responsabilidade social e sustentabilidade**. São Paulo: Atlas, 2006.
- DUAN, J.; SELKER, J.; GRANT, G.E. Evaluation of Probability Density Functions in Precipitation Models for the Pacific Northwest. **Journal Of The American Water Resources Association**. v.34, n.3, p617-627, 1998.
- FISCH, G. Distribuição da precipitação em Taubaté, Vale do Paraíba, SP. **Revista Biociências, Taubaté**, v. 5, n. 2, p. 7-11, 1999.
- IBGE CIDADES, RONDÔNIA, Disponível em: <http://www.cidades.ibge.gov.br/xtras/perfil.php?lang=&codmun=110004&search=||infogr%E1ficos:-informa%E7%F5escompletas>>. Acesso em 13 de janeiro de 2016.
- LEÃO, V.G. **Água tratada: formação de trihalometanos pelo uso do cloro e os riscos potenciais à saúde pública em cidades da mesorregião do leste rondoniense**, Dissertação de Mestrado, Mestrado em Ciências da Saúde, Universidade de Brasília, DF, 131p . 2008.
- METEOBLUE, Precipitação. Disponível em: <https://content.meteoblue.com/pt/ajuda/parametros/precipitacao>>. Acesso em 02 de janeiro de 2016.

MORETTIN, P. A.; BUSSAB, W.O. **Estatística básica**. São Paulo: Saraiva, 2004.

NAGHETTINI, M.; PINTO, E. J.A. **Hidrologia Estatística**. Serviço Geológico do Brasil - CPRM, Belo Horizonte, MG, 600p, 2007. Disponível em:
<http://www.cprm.gov.br/publique/media/livro_hidro_estatistica.zip> Acesso em 10 de dezembro de 2015.

NETO, D. D.; ASSIS, J.; TIMM, L. C.; MANFRON, P. A.; SPAROVEK, G.; MARTIN, T. N. Ajuste de modelos de distribuição de probabilidade a séries históricas de precipitação pluvial diária em Piracicaba–SP. *Rev. Bras. Agrometeorologia*, v. 13, n. 2, p. 273-283, 2005.

NIMER, E. Climatologia do Brasil, vol. 4. **SUPREN/IBGE**, 1979.

PEARSON, K. On the dissection of asymmetrical frequency curves. **Phil. Trans. Roy. Soc**, v. 185, n. pt 1, p. 71-110, 1894.

PLACKETT, R.L. Karl Pearson and the chi-squared test. **International Statistical Review/Revue Internationale de Statistique**, p. 59-72, 1983.

PORTAL BRASIL, CONHEÇA OS BIOMAS BRASILEIROS. Disponível em:
<<http://www.brasil.gov.br/meio-ambiente/2009/10/biomas-brasileiros>> Acesso em 15 de abril de 2016.

R Development Core Team (2006). R: **A Language and Environment for Statistical Computing**, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>, 2006.

SILVA, J. C.; HELDWEIN, A. B.; MARTINS, F. B.; TRENTIN, G.; GRIMM, E. L. Análise de distribuição de chuva para Santa Maria, RS. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 11, n.1, p.67-72, 2007.

THOM, H.C.S. A note on the gamma distribution. **Monthly Weather Review**, v. 86, n. 4, p. 117-122, 1958.

APENDICE – ROTINA DA ANÁLISE DO AJUSTAMENTO

```
rm(list=ls())#Apagar todas as variáveis que estão gravadas no R #
ls()
```

```
#### Script do R para a análise de ajustamento ####
```

```
#### Alterar o diretório para a inserção dos dados ####
getwd()
setwd("D:/Desenvolvimento do TCC")
getwd()
```

```
#### Ou utilizar comando attach para inserir os dados ####
dadosajuste #
attach(dadosajuste)
is.vector(dadosajuste)
h = as.vector(dadosajuste) # Transformando em vetor
h
```

```
#### Pacotes a serem utilizados no ajustamento ####
require(MASS)
library(riskDistributions)
```

```
#### Realizando a análise pelo pacote MASS ####
## Ajuste das distribuições Normal, Exponencial, Log-Normal, Weibull utilizando o teste de
ks ##
```

```
help(fitdistr)
```

```
## Ajuste da Normal ##
fitdistr(h$x, "normal")
ks.test(h$x, "pnorm", mean=mean(h$x), sd=sd(h$x))
```

```
## Ajuste da Exponencial ##
fitdistr(h$x, "exponential")
ks.test(h$x, "pexp", rate = 0.0035047138)
```

```
## Ajuste da Log-Normal ##
fitdistr(h$x, "LogNormal")
ks.test(h$x, "plnorm", meanlog = 5.56001858, sdlog = 0.46425765)
```

```
## Ajuste da Weibull ##
fitdistr(h$x, "Weibull")
ks.test(h$x, "pweibull", shape = 2.9906105, scale = 320.4357338)
```

```
#### Realizando a análise pelo pacote riskDistributions ####
## Pacote que realiza análise das distribuições, e apresenta os gráficos ##
res1<-fit.cont(h$x)
```